DOCUMENT RESUME

ED 349 330 TM 018 960

AUTHOR Hearne, Jill; Schuman, Steven
TITLE Portfolio Assessment: Implementation and Use at an Elementary Level.
PUB DATE 92
NOTE 69p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Academic Achievement; *Educational Assessment; Elementary Education; Elementary School Students; Elementary School Teachers; Evaluation Methods; *Portfolios (Background Materials); Questionnaires; Research Utilization; School Demography; Student Adjustment; *Student Evaluation; *Teacher Role
IDENTIFIERS *Performance Based Evaluation; *Washington

ABSTRACT

Exploratory work in the implementation of a portfolio assessment system at an elementary school in Washington State is described. The effectiveness of the assessment strategy is examined in relation to: (1) teacher implementation levels; (2) student demographic variables; and (3) teacher ratings of students' academic and social adjustment levels plus portfolio use. Students (N=324) in grades 1 through 5 completed a year-long task of selecting writing and mathematics work samples for their portfolios. Structured and free response measures of self-reflection were gathered on the portfolios' contents. The district's records provided student demographic information. Students were tested using the Iowa Tests of Basic Skills and were rated by their teachers, who also rated the portfolios. Teachers completed a questionnaire to identify their concerns about knowledge and use of portfolio information. The level of implementation by staff was higher than might have been predicted given the recent introduction of portfolios. Portfolios differentiated among demographic factors in the same ways as did other assessment measures, suggesting that they do not provide a more unfair picture than other measures. Evidence to date supports the utility of portfolios in an elementary school setting. Six tables present study data, and an appendix containing 49 tables and constituting about two-thirds of the document provides supplemental information about student performance. There is an 18-item list of references. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Portfolio Assessment: Implementation and Use at an Elementary Level

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Jill Hearne, Principal
Steven Schuman, Director of Program Evaluation
Federal Way School District
Federal Way, Washington

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
JILL HEARNE
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED349380

A. Objectives

This paper describes exploratory work in the implementation of portfolio assessment at an elementary school in Washington State. After one year of a schoolwide portfolio assessment project, it examines the effectiveness of the assessment strategy in relation to (1) teacher implementation levels, (2) student demographic variables, and (3) teacher ratings of students' academic and social adjustment levels plus portfolio usage.

B. Perspective

Local accountability and national goals may seem to be dichotomous forces shaping educational policies for the '90's. On every front, there are calls for "proof" of local, state, and national efficacy in meeting both the nationally established goals and a myriad number of locally established goals, many of which can be conflicting or arbitrary.

Many traditional normative, achievement measures (e.g., SAT, ITBS, MAT, CTBS) have been examined and found wanting. Critics (Mitchell, 1992; Resnick, & Resnick, 1985; Shepard, 1991;Wolf, Bixby, Glenn, & Gardner, 1991) claim that the multiple-choice tests are culturally biased, irrelevant, poorly aligned, and/or measure the wrong things (lower order versus higher order thinking). Moreover, they appear to teach skills, such as guessing and quick response recognition finding, that are inconsistent with learning and demonstrating sound problem solving skills. Many educators today maintain that the use of short answer or multiple-choice tests runs counter to the development of those critical thinking, problem solving strategies so necessary in a changing world (Resnick & Resnick, 1989).

Increasing favor is being given to the conception of "authentic assessment" or "performance assessment", wherein the student is asked to think and use information in some task that resembles real life activity. One example is writing an essay; another is balancing a checkbook. Proponents argue that real life demonstrations would be free (or at least more free) of cultural bias, encourage the teaching of creative thought processes and serve as a more accurate (i.e., valid) index of a student's proficiency in a given content or concept area (McLean, 1990; Wiggins, 1989).

Opponents maintain that it is difficult to aggregate authentic assessment beyond a classroom to a building, district or state level and that evaluation of such tasks would be prohibitively expensive, less reliable, more open to cheating and less generalizable

TM018960

BEST COPY AVAILABLE

(Beck, 1991; Fremer, 1991; Mehrens, 1991; Williams, Phillips, & Yen, 1991). Mehrens (1992) warns educators to go slow and "...not jump on any "performance assessment for accountability" bandwagon before [cognitive psychology] theories are much more thoroughly understood [as support for the use of performance assessment]."

Nonetheless, several states, notably Arizona, Kentucky, Oregon and Vermont, have undertaken to develop writing and other performance-based assessments that are more authentic to future student life-role activities (Vermont, 1990). Conventional testing with its uniform administration, empirically developed test items, standardized and operationalized scoring procedures, efficient processing, and norming of tests for appropriate comparability met the needs of those who needed such information to verify student assignments, allocate resources and document accountability. While it provided the lay person and school district patrons with some feeling of objectivity, many, both in and out of education, now feel that the needs of the educator and the student are not being met in meaningful ways by conventional testing. No more critical purpose than this underlies the changes asked for in the assessment movement of the 1990's. Conventional assessments have not promoted the student's participation in acquiring self-assessment skills, of enjoying the challenge of purposeful inquiry into real life problems, and of coming up with unique methods of showing attained skills for expected standards of performance. (Schuman, 1992).

In order to better understand the effectiveness of these reforms, four questions must be answered. First, what kinds of assessments should be administered, and to whom? Second, how many assessments should be given at the national level? at the local level? Third, what are we assessing for? What do we hope to know about the child during the assessment, and what can we gain from such assessments? And fourth, how do the characteristics and results of these alternative assessments compare to those of the more conventional measures? This paper hopes to provide useful data to guide those seeking answers to these questions of educational accountability.

This paper will describe and investigate the following questions of the continuing portfolio project:

- What is the level of implementation by school staff of the key features in the portfolio assessment project?
- Do disaggregated student demographic characteristics show differences in terms of quality of portfolios and attainment of basic skill achievement?
- Are there differences between teacher ratings of student portfolios, social adjustment or academic skills and conventional test scores?

C. Methods:

The primary vehicle under exploration in this project was the use of student *"Portfolios"* at a K-6 elementary school in a metropolitan school district. The working definition of "portfolio assessment" is that it is a purposeful collection of student work which follows four guidelines.

1. The student participates in the selection of portfolio contents.

2. The criteria for selection must be evident.

3. The criteria for judging merit must be evident.

4. There must be evidence of the student's self-reflection about a) the merits of the work or the meeting of a standard, b) personal growth in the development of the skills necessary for the work, and/or c) proposed direction for future efforts. (Meyer, Schuman, and Angello, 1990)

Collected student work in the portfolio included samples of on-going classroom work or specific indicators such as attitude surveys, number of books read, checklists, etc. Over the year, each student in grades 1-6 constructed with teacher guidance such a collection.

D. Data Source

Students in grades kindergarten through six completed a year long task of selecting writing and math work samples for their portfolio. In some cases additional work samples were included from penmanship, art, and reading. Structured and free response measures of self-reflection were gathered on content selected for the portfolio. Culminating activities at the end of the year included the selection of work exemplars for the next year's teacher and for the student's family.

Information from the district's computerized data base provided student demographic characteristics for analysis. Included for this study were the areas of gender; race; special program involvement; and socioeconomic status, as measured through participation in the free and reduced fee lunch program.

Students were tested in late April with the Iowa Test of Basic Skills, Form G (Riverside Publishers, 1985). This was a regularly scheduled assessment occurred each year for all students in grades 1 through 6. Their normal curve equivalent (NCE) scores in reading comprehension, language arts total, math total and total battery were used in this study as measures of conventional assessment.

Grade K-5 teachers as part of an annual task, rated the students on their academic and social readiness for the next grade. Teachers' ratings were scaled from 1 to 5 with 1 being "excellent" and 5 being "low." (See the Appendix for a copy of the scale.) The ratings from the five point scales were part of a standard procedure completed in June of each year to balance classes heterogeneously for the next year by random computer sorting of the ratings.

To rate quality of the portfolios, these same teachers used a new five point scale. The scale attempted to holistically gage the entire portfolio's completeness in meeting the four guidelines for portfolios mentioned above. A portfolio that had work samples and supporting documentation which told a clear story of a student's skill progress, showed evidence of self-reflection, and included insights into the student beyond the work

samples, themselves, was rated as an outstanding portfolio, i.e., a rating of 1. At the other end of the scale an unacceptable portfolio (rated as 5) did not have minimal elements of the guidelines, little or no self-reflection, and showed no evidence of how or why samples were selected to show progress. A copy of this scale is also in the Appendix.

Teachers' ratings of their students' portfolios were used as an estimate of authentic assessment. This is because the first year action plan of the project did not include external evaluative measures of the portfolio's contents or the portfolio as a whole.

Finally to assess teacher level of implementation of the portfolio project, each teacher was administered the Stages of Concern Questionnaire (CBAM, 1974; Hall, George, and Rutherford, 1979) in June. This 35 item scale delineated teachers' concerns about their knowledge and use of the portfolio innovation. Teachers with greater concerns about personal and informational issues were rated as being lower implementers of the portfolio. But for those teachers who had lower concerns in these areas and higher concerns about the consequences of the portfolio on their students or fellow colleagues, these teachers were rated as having concerns that followed from higher levels of implementation. This study provides baseline data for an ongoing investigation of innovation adaptation over time.

In summary, 324 students in classrooms of grades one through five were used in this study. These classrooms had student level data from both teacher ratings and/or achievement tests. 120 students were not included because of nontesting in achievement (all kindergartners) and no teacher ratings (all sixth graders). For the implementation data, seventeen teachers in grades kindergarten through grade six supplied questionnaire feedback (100% return rate).

E. Results

STATVIEW 2.0 (Abacus Concepts, 1987) was used to statistically analyze the relationships among portfolios, student outcomes as measured by standardized tests and teacher rankings, the teacher implementation variable, and student demographic characteristics.

Portfolio Implementation

The first question of investigation concerned the level of implementation of key features of the innovation (i.e., portfolio assessment) by the school staff. To estimate implementation levels, the staff rated their adoption of portfolio assessment tenets, as described in the definition above on the Stages of Concern (SOC) Questionnaire.

Analysis of the questionnaire information for all the 17 teachers revealed a profile approximating the expected distribution reported in the literature for early users of an innovation (Hall, Wallace, and Dossett, 1973). See Figure 1. The mean intensity levels for the SOC scales are shown in Table 1. The Awareness scale is an early implementation concern while Refocusing an innovation is a later concern of more intensive implementers. No differences across the scales were found for gender or grade level among the teachers on the SOC Questionnaire. See Tables 2 and 3.

Figure 1

Profile of Stages of Concern for Adelaide Teachers In the Portfolio Project's Initial Phase: Spring, 1991

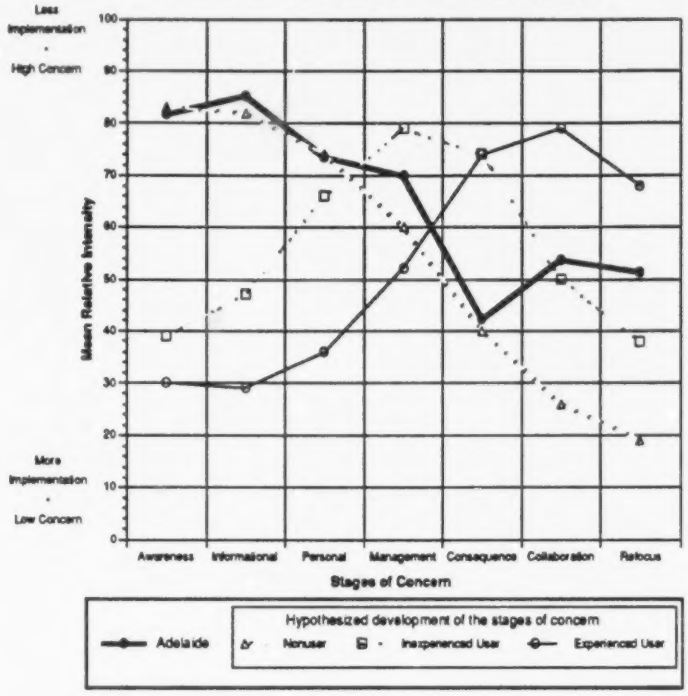


Table 1

Teacher's Self-Report Data on the Stages of Concern Questionnaire

Stages of Concern	Count	Mean	Standard Deviation	Standard Error
Awareness	17	81.65	16.31	3.96
Informational	17	85.18	13.6	3.3
Personal	17	73.59	19.04	4.62
Management	17	70.12	28.83	6.99
Consequence	17	42.29	30.78	7.47
Collaboration	17	53.71	26.12	6.33
Refocus	17	51.35	26.63	6.46

Table 2

Gender Differences in Stages of Concern Data for Adelaide Teachers: Spring, 1991

Stages of Concern		Mean	Standard Deviation	t-score	P value
Awareness	Females (N=10)	74.4	16.85	-2.05	0.0614
	Males (N=5)	91.2	9.35		
Informational	Females (N=10)	81.6	15.17	-1.09	0.3239
	Males (N=5)	89.6	11.91		
Personal	Females (N=10)	66.1	20.65	-1.56	0.1429
	Males (N=5)	81.6	10.5		
Management	Females (N=10)	39.2	30.55	-1.65	0.1224
	Males (N=5)	64.6	21.45		
Consequence	Females (N=10)	35.4	30.04	-1.21	0.2483
	Males (N=5)	56.4	35.2		
Collaboration	Females (N=10)	60.9	27.92	1.34	0.2018
	Males (N=5)	41	24.9		
Refocus	Females (N=10)	53.6	27.61	0.28	0.7862
	Males (N=5)	49.2	31.94		

Table 3

Grade Grouping Differences in Stages of Concern Data for Adelaide Teachers: Spring, 1991

Stages of Concern	Grade Grouping	Mean	Standard Deviation	F-test	P value	Sign. Group Differences
Awareness	K-2	77.25	14.77	1.18	0.338	NS
	3-4	78.3	20.71			
	5-6	95.33	6.66			
Informational	K-2	83.75	13.54	0.23	0.8097	NS
	3-4	83.2	15.55			
	5-6	89.67	8.5			
Personal	K-2	67.62	23.11	0.45	0.6486	NS
	3-4	76.8	17.11			
	5-6	77	5.57			
Management	K-2	61.75	29.44	1.07	0.37	NS
	3-4	66.2	33.95			
	5-6	90	6.24			
Consequence	K-2	37.5	34.16	0.11	0.9003	NS
	3-4	46	34.78			
	5-6	43.33	28.45			
Collaboration	K-2	54.25	25.56	1.87	0.194	NS
	3-4	64.2	31.23			
	5-6	29	1.73			
Refocus	K-2	46.5	27.86	0.35	0.7105	NS
	3-4	58.6	29.16			
	5-6	38	28.51			

A median split was completed on data for each scale and teachers were coded 1 for high implementation-low concern, and 2 for low implementation-high concern. These codes were then summed, resulting in an index score ranging was from 7 to 14. This was done to regroup the teachers into high and low implementers based upon all the SOC scores. Teachers rating themselves as intermediate ("high") users received a score of 7, 8, 9 or 10. Those rating themselves as novice or non-implementing ("low") users received scores of 11, 12 or 13. There were no scores of 14. See Table 4.

It was believed this index would show the likelihood that the higher the level of overall concern (i.e., higher score) would be related to lower levels of implementation and low concern related to higher implementation. Teacher portfolio ratings of their students were, however, not significantly different between the high and low implementers ($t = 0.41$; $p < .68$). "High" implementing teachers (N=8) rated their students no differently on having more highly evolved portfolios than low implementing teachers, ratings of 2.98 and 3.05 respectively. All other teacher ratings and student achievement scores were also not different between high and low implementers.

Self-ratings of being non users or high implementers on a separate (nonSOC) scale produced inconsistent correlations to the SOC Questionnaire data. But known teacher behavior as observed by the teacher's supervisor was compared with the SOC Index by the first author. These Stages of Concern rankings validated this investigator's informal groupings of teachers based on observation (Hearne, 1992). Teachers who were perceived by observation to be high implementers did indeed get the lowest scores on the SOC Index, while those teachers seen to be low or non-implementing had the highest scores. Examination of the Stages of Concern profile (Figure 2) demonstrates

Table 4

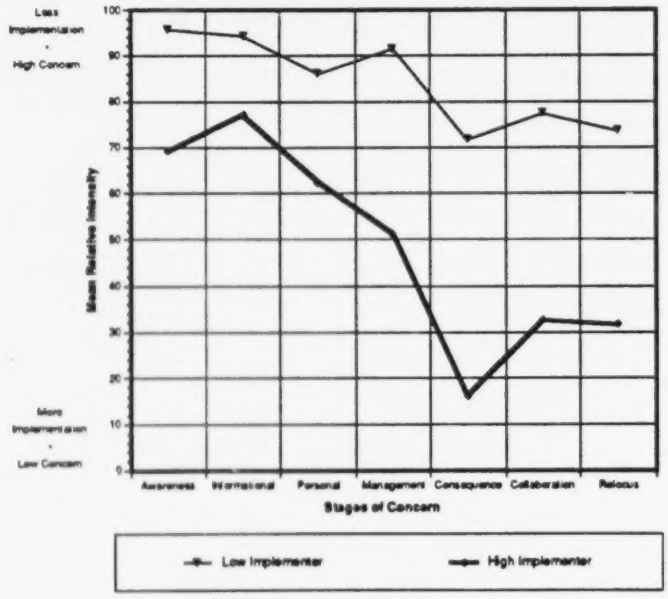
Descriptive Statistics for High and Low Implementing* Teachers on the Stages of Concern Questionnaire

	Count	Low Implementer (N=9)	High Implementer (N=8)	t-score	P value
Awareness	17	95.62	69.22	-5.83	0.0001
Informational	17	94.25	77.11	-3.3	0.0049
Personal	17	86	62.56	-3.17	0.0063
Management	17	91.38	51.22	-3.98	0.0012
Consequence	17	71.75	16.11	-9.79	0.0001
Collaboration	17	77.38	32.67	-7.2	0.0001
Refocus	17	73.62	31.56	-5.4	0.0001

*Based on a median split of the scale data

Figure 2

Stages of Concern Profile for Low and High Implementing Teachers In the Portfolio Project's Initial Phase: Spring, 1991



the differences between the high and low implementer groups on the seven variables.

Thus, self-reported implementation to the portfolio innovation was related to discernible differences among teachers perception of quality portfolios but not for other student outcomes of achievement or teacher perceived ratings of adjustment and ability.

The second question investigated the extent to which student demographic characteristics were reflected in portfolio ratings and other student outcome measures.

Student outcomes examined included: standardized data from the Iowa Test of Basic Skills (ITBS) scores in reading, math, language arts, and total battery, as well as teacher ratings on students' academic ability, social responsibility, and portfolio quality.

Student demographic factors analyzed by these outcome measures were grade level, gender, ethnicity, and participation in the free and reduced fee lunch program (socioeconomic estimate) or the special educational programs of the school, that is special education, Chapter 1/LAP or gifted education programs.

Table 5 summarizes the 49 analyses of these seven outcome measures and seven demographic factors. Analysis of Variance tables for these data are contained in the Appendix.

Student Characteristics

Gender: Only social adjustment ratings and mathematics test scores revealed any significant differences. Girls were rated significantly more socially adjusted (2.56 vs. 1.98; $p < .0001$) by teachers and boys had higher math performance (NCE) scores (64.2 vs. 59.3; $p < .02$). Boy and girl differences on quality of portfolios were not evident. See Tables 1-7 in the Appendix.

Ethnicity: Academic and portfolio ratings plus language arts scores were significantly different among the students from different racial backgrounds. However, some the number of students representing a racial group were quite small. See Tables 8-14.

Socioeconomic Status: Four of the student measures were significantly different for students participating in the school's free or reduced fee lunch program. These students were less ready academically (2.93 vs. 2.37; $p < .002$) or socially (2.57 vs. 2.22; $p < .03$) for their next grade level and lower in their language arts (58.2 vs. 66.4; $p < .01$) and overall (Total Battery: 56.7 vs. 62.5; $p < .05$) achievement scores. No differences were found for quality of portfolios among low and higher socioeconomic students. See Tables 15-21.

School Program Characteristics

Grade level: Significant differences were found in all areas, except social adjustment ratings for next year's placement. Differences show that first grade academic and portfolio ratings are significantly lower than all the other grades but their reading levels are higher. Other achievement scores vary irregularly across the grades. Se Tables 22-28.

Special education: All outcome factors were significantly lower, except for language arts scores. See Tables 29-35. The magnitude of the differences between special education and nonspecial education students was as great in portfolio quality as it was in the other outcome areas.

Table 5

Analysis of Variance Summaries of Teacher Ratings and Tested Achievement for Student and School Factors (Grades 1-5)

	Teacher Ratings			Tested Achievement (Iowa Test of Basic Skills)			
	Academic Rating	Social Adj. Rating	Portfolio Rating	Reading	Language Arts	Mathematics	Total Battery
Student Characteristics							
Gender	Non (p<.08)	Significant (p<.0001)	**Non (p<.2)**	Non (p<.23)	Non (p<.42)	Significant (p<.02)	Non (p<.95)
Ethnicity	Significant (p<.005)	Non (p<.22)	**Significant (p<.03)**	Non (p<.60)	Significant (p<.02)	Non (p<.43)	Non (p<.15)
Socioeconomic Status*	Significant (p<.002)	Significant (p<.03)	**Non (p<.39)**	Non (p<.18)	Significant (p<.02)	Non (p<.08)	Significant (p<.05)
School Programs							
Grade Level Differences	Significant (p<.02)	Non (p<.89)	**Significant (p<.0001)**	Significant (p<.001)	Significant (p<.02)	Significant (p<.0002)	Significant (p<.004)
Special Education	Significant (p<.0001)	Significant (p<.0001)	**Significant (p<.004)**	Significant (p<.0001)	Non (p<.08)	Significant (p<.005)	Significant (p<.0002)
Chapter 1/LAP**	Significant (p<0001)	Significant (p<.0001)	**Significant (p<.006)**	Significant (p<.0001)	Significant (p<.0001)	Significant (p<.0006)	Significant (p<.0001)
Gifted Program	Significant (p<.0001)	Non (p<.17)	**Significant (p<.003)**	Significant (p<.02)	Significant (p<.008)	Significant (p<.003)	Significant (p<.0002)

*Participation in Free & Reduced Fee Lunch Program

**LAP is a state funded Chapter 1 type program

Analysis of Variance Data Tables in Appendix

Chapter 1/LAP: All outcomes were significantly different. Chapter 1 participants had lower teacher ratings and lower performance levels on the achievement test battery. See Tables 36-42.

Gifted education: Social adjustment ratings were not significantly different between gifted program participants and non participants. All other areas found participants to have better teacher ratings and higher levels of achievement. See Tables 43-49.

In summary, these findings show portfolio ratings to differentiate among demographic factors in similar ways that other outcome measures do. Teacher perceptions of student quality in portfolios has some evidence of ethnic differences but not gender or socioeconomic differences and, just like conventional achievement measures, portfolio ratings separate out participants in various school programs.

Correlational Analysis

Table 6 shows the intercorrelation of student and teacher measures. Portfolio rating correlate around .51 with the academic rating and .25 with the social adjustment rating. There is essentially a lack of relationship between the portfolio ratings, ranging around -.16 to +.06, and the conventional measures of student achievement. (Note. The teacher rating scales were inversely related to the test result scales. Therefore, positive teacher ratings appear as negative correlations with the test scores.)

The correlations between the academic rating and the conventional measures was around -.47 to .-.53. This was about as strong as the relationship between the portfolio rating and the academic rating, although the portfolio ratings were not correlated at all with the test scores.

F. Conclusions

The level of implementation by school staff was higher than might have been predicted given the fact that they had been first introduced to portfolios in October, 1990 and a state-wide teacher strike closed schools for the month of April. Awareness, informational, personal, and management concerns were all relatively high, while consequences was lower than one would expect.

One year is a very short time to track a change which requires as great a shift in classroom management practices and instructional focus as does portfolio assessment. We think that our lack of student data at grades kindergarten and six weakened our analysis. This was due to the fact that it was not normal procedure for sixth grade teachers to rank students on social and academic rankings because the data are used for class placement. The investigators will remedy this procedure in the second year of data collection.

On demographic variables such as race, gender, and SES as measured by free and reduced price lunch, there is indication that portfolios may provide a more fair measure of a students' capabilities or at least not reflect a picture any different than that portrayed by conventional measures. Concerns about gender, race and socioeconomic bias are ongoing in critiques regarding evaluation. Portfolio assessment can contribute

Table 6

Correlational Matrix of Student and Teacher Measures (Grades 1-5)

	ITBS* Acheivement Test Results				Student Ratings**			Stages of Concern Teacher Self-Report Scales						
	Reading	Language Arts	Mathematics	Total Battery	Portfolio	Academic	Social	Awareness	Informational	Personal	Management	Consequences	Collaboration	Refocus
Reading														
Language Arts	0.53													
Mathematics	0.53	0.56												
Total Battery	0.77	0.78	0.74											
Portfolio	0.06	-0.18	-0.13	-0.13										
Academic	-0.47	-0.53	-0.51	-0.61	0.51									
Social	-0.35	-0.34	-0.32	-0.37	0.25	0.51								
Awareness	-0.02	-0.01	-0.04	0.03	-0.51	-0.2	-0.05							
Informational	-0.13	0.09	-0.05	-0.01	0.18	0.02	0.03	0.2						
Personal	-0.34	-0.18	-0.01	-0.14	-0.35	-0.09	0.02	0.1	0.25					
Management	-0.23	-0.08	0.03	-0.06	-0.17	-0.02	0.00027	0.35	0.38	0.79				
Consequences	-0.24	-0.05	0.00266	-0.08	-0.1	-0.05	0.04	-0.03	0.44	0.58	0.38			
Collaboration	-0.2	0.01	-0.1	-0.13	-0.18	-0.09	-0.01	-0.32	0.08	0.11	-0.25	0.48		
Refocus	-0.15	0.11	-0.01	-0.03	-0.13	-0.09	-0.05	-0.07	0.45	0.55	0.46	0.64	0.44	
SOC Index	-0.11	0.05	-0.02	-0.01	-0.01	-0.03	-0.03	0.04	0.52	0.63	0.81	0.65	0.34	0.66

*Iowa Test of Basic Skills **Rating scae is inversely related to test results; therefore, higher test scores are negatively correlated to positive teacher ratings.

to school improvement efforts if portfolios can paint a picture of a student's abilities which helps teachers with instructional decision making but does not indicate bias on demographic variables.

The relationships between special education, Chapter I, gifted education and portfolio ratings were in the expected direction. This is the most meaningful part of this investigation in the eyes of the authors. These indicate that with increased utilization of portfolios, teachers will be able to use portfolios for assessment purposes with some assurance that they reflect academic abilities as least as well as measured by standardized tests. Further, if future data continue to show these relationships, perhaps the public and policymakers will come to feel more comfortable regarding the validity of authentic assessment measures compared with these traditional measures.

G. Educational Importance of the Study

Given the little time involved in adaptation of the innovation at the June, 1991 data collection point, we are pleased with the information gathered to date and heartened by the fact that correlations and levels of significance indicate that further investigation is warranted. As the teachers increase in their usage of portfolios we predict that their interests will move away from the management and personal issues to the collaboration and refocussing issues. When this occurs we can also predict a tighter coupling between standardized test scores and teacher ratings of portfolios. We will also continue to pursue the possibility of the portfolio being a more fair gender and ethnically instrument than standardized tests.

The current exploration of portfolio use at Adelaide Elementary School is designed to further our knowledge of portfolio use both in but also beyond the classroom level. Our ability to drive national, state, and local agendas for educational reform lies in our ability to assert technical expertise in using accountability measures that are first driven by qualitative rather than quantitative measures. If we want more students adept at thinking and problem-solving, we need to measure them in ways that have real life validity as well as psychometric validity and reliability. Moreover, teachers and parents are more likely to embrace evaluation techniques which broaden student understanding of what constitutes quality work rather than settle for yearly national testing which teaches students skills for taking a test.

References:

Beck, M. D.. (1991). Authentic assessment for large-scale accountability purposes: Balancing the rhetoric. Paper presented at the annual meeting of the American Educational Research Association. Chicago, Ill.

Fremer, J. (1991). Performance testing and standardized testing: What are their proper places. Paper presented at the annual meeting of the American Educational Research Association. Chicago, Ill.

Hall, G. E., Wallace, R. C.., Jr. & Dossett, W.A. (1973) *A Developmental Conceptualization of the Adoption Process Within Educational Institutions.* Austin: Research and Development Center for Teacher Education, The University of Texas Press.

Hall, G., George, A., Rutherford, W. (1979) *Measuring Stages of Concern About The Innovation.* Austin, Texas: The University of Texas Press.

Hearne, J. (1992), Portfolio Assessment: Tracking its implementation and use in one elementary school. In J. Bamberg (Ed.) *Assessment: How do we know what they know* . Union, Washington: Association for Supervision and Curriculum Development.

McLean, L.D. (1990). Time to replace the classroom test with authentic measurement. *Albert Journal of Educational Research, 36,* 78-84.

Mehrens, W.A. (1991). Using performance assessment for accountability purposes: Some problems. Paper presented at the annual meeting of the American Educational Research Association. Chicago, Ill.

Mehrens, W.A. (1992). Using performance assessment for accountability purposes. *Educational Measurement: Issues and Practices, 11(1),*, 3-9, 20.

Mitchell, R. (1992). *Testing for Learning: How new approaches to evaluation can improve American schools.* New York, NY: The Free Press.

Meyer, C., Schuman, S., and Angello, N. (1990). *NWEA White Paper On Aggregating Portfolio Data.* Lake Oswego, Oregon: Northwest Evaluation Association.

Resnick, L. B., & Resnick, D. P. (1985). Standards, Curriculum and Performance: A historical and comparative perspective. *Educational Researcher, 14,* 5-21.

Resnick, L. B., & Resnick, D. P. (1989). Assessing the thinking curriculum: New tools for educational reform, In B. R. Gifford and M. C. O'Connor (Eds.) *Changing Assessments: Alternative views of aptitude, achievement and instruction.* Boston: Kluwer Academic Publishers.

Schuman, S. (1992). Alternative Assessment: An unfolding story. In J. Bamberg (Ed.) *Assessment: How do we know what they know* . Union, Washington: Association for Supervision and Curriculum Development.

Shepard, L. A. (1991). Psychometrician's beliefs about learning. *Educational Researcher, 20* (7), 2-16.

Vermont State Department of Education. (1990). Vermont Mathematics Portfolios. Montpelier, VT: 05602.

Wiggins, G. (1989). A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan*, May, 703-712.

Williams, P.L., Phillips, G.W., & Yen, W.M. (1991). Measurement issues in high stakes performance assessment. Paper presented at the annual meeting of the American Educational Research Association. Chicago, Ill.

Wolf, D., Bixby, J., Glenn, J., & Gardner, H. (1991). To use their minds well: Investigating new forms of student assessment. In G. Grant (Ed.), *Review of research in education* (pp. 312-74). Washington, DC: American Educational Research Association.

Appendix

SOCIAL AND ACADEMIC SCALES FOR CLASS PLACEMENT

SOCIAL

1. The student is social and emotionally adjusted and is never a problem in class. This student is more likely to be a daily asset by being helpful. Completely independent. Never needs help.

2. The student is well-adjusted and is rarely a problem. The student is usually helpful. He/she is quite independent and needs help once in a while.

3. The student is average in social and emotional adjustment. He/she has problems with behavior once in a while but it is usually not serious. Independence level is average--they need some help on a daily basis.

4. The student is in need of social and emotional adjustment. This student may be referred to the counselor and/or MDT team. There are daily behavior problems and sometimes they are serious. They are not considered independent learners.

5. The student is socially and emotionally disabled. There are serious, daily behavior problems both in the room and on the grounds and the student is referred to the counselor and principal. This student is school dependent for academics and socialization.

ACADEMIC

1. The student is an A student and is at the head of the class in all or most academic respects. They are curious and willing to do extra work without pressure. Want good grades. Excellent homework.

2. The student is a B student with more than one A. They are interested in good grades and want to do well. Try hard. Does extra work much of the time. Does all homework.

3. The student gets C's with the occasional B or D. Average in effort, interest and accomplishment. Participates in extra work once in a while. Probably excels in one area such as art or handwriting. Average homework completion rate.

4. The student is below average in both grades achieved and effort expended. Often seems to be not interested. Does not do much homework and is careless with the work done. This student may have difficulty accepting responsibility. This student may receive special help from support staff.

5. The student gets many failing grades in spite of support staff assistance. He/she is below grade level in more than one academic area. Attitude is poor and homework is seldom done. Frequent re-teaching is necessary. Is probably not able to do the work.

❊ Portfolio Rating Scale ❊

Steven Schuman

Directions: Below is a definition of portfolios and a rating scale for use in refereeing the quality of the materials comprising the portfolios in your classroom. Please use the definition and the scale to rate each child on the quality of their portfolio as a whole, not the separate elements used to make up their portfolio.

Place a value (1 to 5) from the rating scale on the computer sheet previously used to rate academic and social skills. Return the completed computer sheet to your principal.

•Definition of a portfolio standard.

A portfolio is a (1) purposeful collection of student work that shows student (2) effort, progress, or achievement in one or more areas. The collection includes evidence of (3) student self-reflection and (4) participation in setting the focus, selecting contents, and judging merit. Activities are (5) guided by performance standards. A portfolio communicates what is learned and why it is important.

•Holistic rating scale that should reflect the sum of the portfolio parts.

1	**Outstanding**	Exemplary, defines the highest standards for portfolio assessment.
3	**Acceptable**	Satisfies the definition of a portfolio standard.
5	**Unacceptable**	Does not meet the minimal standards reflected in the definition.

1 Outstanding Portfolios tell a *clear story* of a unique student and his/her context. The portfolio user sees evidence of progress/lack of progress via student (and sometimes teacher) selected examples and stated standards. Necessary examples are present to make an informed judgement of status and growth. Self-reflection records are always present and aligned with examples, are insightful, and accurately reflect the student's view of status and direction for growth/change.

2 Better Than Acceptable Portfolios

3 Acceptable Portfolios conform to the definition of a portfolio. They tell a *usable story* or create a picture of a student. By use of the portfolio content, the portfolio user knows why an example(s) was selected for what issue(s). Most examples include self-reflection records from personal points of view that usually relate to an established performance standard(s).

4 Less Than Acceptable Portfolios

5 Unacceptable Portfolios do not meet the minimal elements of the portfolio definition. They tell a *vague or incomplete story* of a student out of context. The portfolio user has no evidence of what instructional issue(s) was addressed and why an example(s) was selected. Inferring what issue(s) and why an example(s) was selected and what standard(s) was used to judge merit dominates the reviewer's process. Little or no record of reflection is evident.

6/19/91

© 1991 Federal Way School District

Table 1

Analysis of Variance Table for Gender by Academic Rating

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	3.67	3.67	3.27	0 .0716
Within groups	329	369.93	1.12		
Total	330	373.6			
Group:	Count	Mean	Standard Deviation	Standard Error	
Males	160	2.55	1.07	0.08	
Females	171	2.34	1.05	0.08	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Males vs. Females	0.21	0.23	3.27	1.81	

Table 2

Analysis of Variance Table for Gender by Social Adjustment Rating

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	27.93	27.93	32.26	0.0001
Within groups	322	278.77	0.87		
Total	323	306.7			
Group:	Count	Mean	Standard Deviation	Standard Error	
Males	158	2.56	0.97	0.08	
Females	166	1.98	0.89	0.07	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Males vs. Females	0.59	0.2*	32.26*	5.68	

*Significant at 95%

Table 3

Analysis of Variance Table for Gender by Portfolio Rating

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	2.96	2.96	1.71	0 .1926
Within groups	259	450.05	1.74		
Total	260	453.02			
Group:	Count	Mean	Standard Deviation	Standard Error	
Males	129	3.05	1.24	0.11	
Females	132	2.83	1.39	0.12	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Males vs. Females	0.21	0.32	1.71	1.31	

Table 4

Analysis of Variance Table for Gender by Reading Comprehension Achievement

Source:	DF:	Sum Squares:	Mean Square:	F-test:	P value
Between groups	1	458.62	458.62	1.46	0.2273
Within groups	318	99660.3	313.4		
Total	319	100118.92			
Group:	Count:	Mean	Standard Deviation	Standard Error	
Males	158	61.77	18.99	1.51	
Females	162	64.17	16.35	1.28	
Comparison:	Mean Difference	Fisher PLSD:	Scheffe F-test:	Dunnett t:	
Males vs. Females	-2.39	3.89	1.46	1.21	

Table 5

Analysis of Variance Table for Gender by Language Arts Achievement

Source:	DF:	Sum Squares:	Mean Square:	F-test:	P value
Between groups	1	235.21	235.21	0.65	0 .4197
Within groups	318	114588.68	360.34		
Total	319	114823.89			
Group:	Count:	Mean	Standard Deviation	Standard Error	
Males	158	64.55	18.26	1.45	
Females	162	66.27	19.66	1.54	
Comparison:	Mean Difference	Fisher PLSD:	Scheffe F-test:	Dunnett t:	
Males vs. Females	-1.71	4.18	0.65	0.81	

Table 6

Analysis of Variance Table for Gender by Mathematics Achievement

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	1953.01	1953.01	5.44	0 .0203
Within groups	322	115671.09	359.23		
Total	323	117624.1			
Group:	Count	Mean	Standard Deviation	Standard Error	
Males	161	64.2	19.61	1.55	
Females	163	59.29	18.28	1.43	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Males vs. Females	4.91	4.14*	5.44*	2.33	

*Significant at 95%

Table 7

Analysis of Variance Table for Gender by Total Battery Achievement

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	1.41	1.41	0.01	0 .943
Within groups	317	87006.68	274.47		
Total	318	87008.09			
Group:	Count	Mean	Standard Deviation	Standard Error	
Males	159	61.75	17.45	1.38	
Females	160	61.88	15.64	1.24	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Males vs. Females	-0.13	3.65	0.01	0.07	

Table 8

Analysis of Variance Table for Racial Background by Academic Rating

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	4	16.7	4.18	3.81	0 .0048
Within groups	324	354.62	1.09		
Total	328	371.32			
Group:	Count	Mean	Standard Deviation	Standard Error	
Hispanic	10	2.8	1.14	0.36	
White	284	2.45	1.06	0.06	
Asian	19	1.74	0.73	0.17	
African American	14	3.07	1.07	0.29	
Native American	2	2	1.41	1	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Hispanic vs. White	0.35	0.66	0.26	1.03	
Hispanic vs. Asian	1.06	0.8*	1.69	2.6	
Hispanic vs. Afr. Am.	-0.27	0.85	0.1	0.63	
Hispanic vs. Nat. Am.	0.8	1.59	0.24	0.99	
White vs. Asian	0.72	0.49*	2.09	2.89	
White vs. Afr. Am.	-0.62	0.56*	1.16	2.15	
White vs. Nat. Am.	0.45	1.46	0.09	0.61	
Asian vs. Afr. Am.	-1.33	0.73*	3.28*	3.62	
Asian vs. Nat. Am.	-0.26	1.53	0.03	0.34	
Africian Am. vs. Nat. Am.	1.07	1.56	0.46	1.35	

*Significant at 95%

Table 9

Analysis of Variance Table for Racial Background by Social Adjustment Rating

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	4	5.5	1.38	1.45	0 .218
Within groups	317	301.06	0.95		
Total	321	306.56			
Group:	Count	Mean	Standard Deviation	Standard Error	
Hispanic	11	2.55	0.93	0.28	
White	277	2.25	0.98	0.06	
Asian	19	1.95	0.78	0.18	
African American	13	2.62	1.12	0.31	
Native American	2	3	0	0	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Hispanic vs. White	0.29	0.59	0.24	0.98	
Hispanic vs. Asian	0.6	0.73	0.66	1.62	
Hispanic vs. Afr. Am.	-0.07	0.79	0.01	0.18	
Hispanic vs. Nat. Am.	-0.45	1.47	0.09	0.61	
White vs. Asian	0.31	0.45	0.44	1.32	
White vs. Afr. Am.	-0.36	0.54	0.43	1.31	
White vs. Nat. Am.	-0.75	1.36	0.29	1.08	
Asian vs. Afr. Am.	-0.67	0.69	0.91	1.9	
Asian vs. Nat. Am.	-1.05	1.43	0.53	1.45	
Africian Am. vs. Nat. Am.	-0.38	1.46	0.07	0.52	

Table 10

Analysis of Variance Table for Racial Background by Portfolio Rating

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	4	19.49	4.87	2.88	0.0234
Within groups	256	433.53	1.69		
Total	260	453.02			
Group:	Count	Mean	Standard Deviation	Standard Error	
Hispanic	8	3.62	1.41	0.5	
White	223	2.93	1.33	0.09	
Asian	16	2.25	0.93	0.23	
African American	12	3.67	1.07	0.31	
Native American	2	2	0	0	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Hispanic vs. White	0.69	0.92	0.55	1.48	
Hispanic vs. Asian	1.38	1.11*	1.49	2.44	
Hispanic vs. Afr. Am.	-0.04	1.17	0.00123	0.07	
Hispanic vs. Nat. Am.	1.62	2.03	0.62	1.58	
White vs. Asian	0.68	0.66*	1.03	2.03	
White vs. Afr. Am.	-0.73	0.76	0.91	1.9	
White vs. Nat. Am.	0.93	1.82	0.25	1.01	
Asian vs. Afr. Am.	-1.42	0.98*	2.03	2.85	
Asian vs. Nat. Am.	0.25	1.92	0.02	0.26	
Africian Am. vs. Nat. Am.	1.67	1.96	0.7	1.68	

*Significant at 95%

Table 11

Analysis of Variance Table for Racial Background by Reading Comprehension Achievement

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	4	933.39	233.35	0.74	0 .5661
Within groups	314	99181.45	315.86		
Total	318	100114.85			
Group:	Count	Mean	Standard Deviation	Standard Error	
Hispanic	10	63.9	16.42	5.19	
White	281	63.12	17.95	1.07	
Asian	18	63.89	12.12	2.86	
African American	9	54.11	22.52	7.51	
Native American	1	77			
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Hispanic vs. White	0.78	11.25	0.00464	0.14	
Hispanic vs. Asian	0.01	13.79	0.000000628	0.00159	
Hispanic vs. Afr. Am.	9.79	16.07	0.36	1.2	
Hispanic vs. Nat. Am.	-13.1	36.68	0.12	0.7	
White vs. Asian	-0.77	8.5	0.01	0.18	
White vs. Afr. Am.	9.01	11.84	0.56	1.5	
White vs. Nat. Am.	-13.88	35.03	0.15	0.78	
Asian vs. Afr. Am.	9.78	14.28	0.45	1.35	
Asian vs. Nat. Am.	-13.11	35.93	0.13	0.72	
Africian Am. vs. Nat. Am.	-22.89	36.86	0.37	1.22	

Table 12

Analysis of Variance Table for Racial Background by Language Arts Achievement

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	4	4469.3	1117.33	3.18	0 .014
Within groups	314	110347.9	351.43		
Total	318	114817.2			
Group:	Count	Mean	Standard Deviation	Standard Error	
Hispanic	10	64.4	17.3	5.47	
White	281	65.81	18.82	1.12	
Asian	18	70.39	14.43	3.4	
African American	9	44.44	24.84	8.28	
Native American	1	62			
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Hispanic vs. White	-1.41	11.87	0.01	0.23	
Hispanic vs. Asian	-5.99	14.55	0.16	0.81	
Hispanic vs. Afr. Am.	19.96	16.95*	1.34	2.32	
Hispanic vs. Nat. Am.	2.4	38.69	0.00373	0.12	
White vs. Asian	-4.58	8.97	0.25	1	
White vs. Afr. Am.	21.37	12.49*	2.83*	3.37	
White vs. Nat. Am.	3.81	36.95	0.01	0.2	
Asian vs. Afr. Am.	25.94	15.06*	2.87*	3.39	
Asian vs. Nat. Am.	8.39	37.9	0.05	0.44	
Africian Am. vs. Nat. Am.	-17.56	38.88	0.2	0.89	

*Significant at 95%

Table 13

Analysis of Variance Table for Racial Background by Mathematics Achievement

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	4	1424.03	356.01	0.97	0 .4214
Within groups	318	116131.44	365.19		
Total	322	117555.47			
Group:	Count	Mean	Standard Deviation	Standard Error	
Hispanic	10	60.9	12.46	3.94	
White	284	61.46	19.45	1.15	
Asian	18	68.44	16.55	3.9	
African American	10	55.8	18.31	5.79	
Native American	1	77			
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Hispanic vs. White	-0.56	12.1	0.00206	0.09	
Hispanic vs. Asian	-7.54	14.83	0.25	1	
Hispanic vs. Afr. Am.	5.1	16.82	0.09	0.6	
Hispanic vs. Nat. Am.	-16.1	39.44	0.16	0.8	
White vs. Asian	-6.99	9.14	0.57	1.5	
White vs. Afr. Am.	5.66	12.1	0.21	0.92	
White vs. Nat. Am.	-15.54	37.67	0.16	0.81	
Asian vs. Afr. Am.	12.64	14.83	0.7	1.68	
Asian vs. Nat. Am.	-8.56	38.63	0.05	0.44	
Africian Am. vs. Nat. Am.	-21.2	39.44	0.28	1.06	

Table 14

Analysis of Variance Table for Racial Background by Total Battery Achievement

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	4	1945.9	486.47	1.79	0 .13
Within groups	313	84913.25	271.29		
Total	317	86859.15			
Group:	Count	Mean	Standard Deviation	Standard Error	
Hispanic	10	57.1	13.13	4.15	
White	280	62.07	16.67	1	
Asian	18	65.89	14.06	3.31	
African American	9	49.33	17.58	5.86	
Native American	1	64			
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Hispanic vs. White	-4.97	10.43	0.22	0.94	
Hispanic vs. Asian	-8.79	12.78	0.46	1.35	
Hispanic vs. Afr. Am.	7.77	14.89	0.26	1.03	
Hispanic vs. Nat. Am.	-6.9	33.99	0.04	0.4	
White vs. Asian	-3.82	7.88	0.23	0.95	
White vs. Afr. Am.	12.74	10.98*	1.3	2.28	
White vs. Nat. Am.	-1.93	32.47	0.00342	0.12	
Asian vs. Afr. Am.	16.56	13.23*	1.52	2.46	
Asian vs. Nat. Am.	1.89	33.3	0.00311	0.11	
Africian Am. vs. Nat. Am.	-14.67	34.16	0.18	0.84	

*Significant at 95%

Table 15

Analysis of Variance Table for Socioeconomic Status by Academic Rating

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	11.04	11.04	10.02	0 .0017
Within groups	329	362.56	1.1		
Total	330	373.6			
Group:	Count	Mean	Standard Deviation	Standard Error	
Participants	41	2.93	1.25	0.2	
NonParticipants	290	2.37	1.02	0.06	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Part vs. NonPart.	0.55	0.34*	10.02*	3.17	

*Significant at 95%

Table 16

Analysis of Variance Table for Socioeconomic Status by Social Adjustment rating

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	4.61	4.61	4.91	0 .0273
Within groups	322	302.09	0.94		
Total	323	306.7			
Group:	Count	Mean	Standard Deviation	Standard Error	
Participants	42	2.57	1.19	0.18	
NonParticipants	282	2.22	0.93	0.06	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Part vs. NonPart.	0.36	0.32*	4.91*	2.22	

*Significant at 95%

Table 17

Analysis of Variance Table for Socioeconomic Status by Portolio Rating

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	1.3	1.3	0.74	0 .3897
Within groups	259	451.72	1.74		
Total	260	453.02			
Group:	Count	Mean	Standard Deviation	Standard Error	
Participants	29	3.14	1.33	0.25	
NonParticipants	232	2.91	1.32	0.09	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Part vs. NonPart.	0.22	0.51	0.74	0.86	

Table 18

Analysis of Variance Table for Socioeconomic Status by Reading Comprehension Achievement

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	568.76	568.76	1.82	0 .1787
Within groups	318	99550.16	313.05		
Total	319	100118.92			
Group:	Count	Mean	Standard Deviation	Standard Error	
Participants	37	59.3	21.63	3.56	
NonParticipants	283	63.47	17.13	1.02	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Part vs. NonPart.	-4.17	6.09	1.82	1.35	

Table 19

Analysis of Variance Table for Socioeconomic Status by Language Arts Achievement

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	2170.38	2170.38	6.13	0 .0138
Within groups	318	112653.51	354.26		
Total	319	114823.89			
Group:	Count	Mean	Standard Deviation	Standard Error	
Participants	37	58.22	23.24	3.82	
NonParticipants	283	66.36	18.18	1.08	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Part vs. NonPart.	-8.14	6.47*	6.13*	2.48	

*Significant at 95%

Table 20

Analysis of Variance Table for Socioeconomic Status by Mathematics Achievement

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	1129.88	1129.88	3.12	0 .0781
Within groups	322	116494.22	361.78		
Total	323	117624.1			
Group:	Count	Mean	Standard Deviation	Standard Error	
Participants	38	56.61	21.04	3.41	
NonParticipants	286	62.41	18.74	1.11	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Part vs. NonPart.	-5.8	6.46	3.12	1.77	

Table 21

Analysis of Variance Table for Socioeconomic Status by Total Battery Achievement

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	1064.02	1064.02	3.92	0 .0484
Within groups	317	85944.07	271.12		
Total	318	87008.09			
Group:	Count	Mean	Standard Deviation	Standard Error	
Participants	36	56.69	19.02	3.17	
NonParticipants	283	62.47	16.12	0.96	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Part vs. NonPart.	-5.77	5.73*	3.92*	1.98	

*Significant at 95%

Table 22

Analysis of Variance Table for Grade Level by Academic Rating

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	4	13.66	3.42	3.09	0 .016
Within groups	326	359.94	1.1		
Total	330	373.6			
Group:	Count	Mean	Standard Deviation	Standard Error	
Grade 1	72	2.79	0.87	0.1	
Grade 2	80	2.4	1.04	0.12	
Grade 3	74	2.41	1.1	0.13	
Grade 4	39	2.13	1.03	0.17	
Grade 5	66	2.33	1.19	0.15	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
1 vs 2	0.39	0.34*	1.32	2.29	
1 vs 3	0.39	0.34*	1.23	2.22	
1 vs 4	0.66	0.41*	2.52*	3.18	
1 vs 5	0.46	0.35*	1.64	2.56	
2 vs 3	-0.01	0.33	0.000254	0.03	
2 vs 4	0.27	0.4	0.44	1.32	
2 vs 5	0.07	0.34	0.04	0.38	
3 vs 4	0.28	0.41	0.44	1.33	
3 vs 5	0.07	0.35	0.04	0.41	
4 vs 5	-0.21	0.42	0.23	0.97	

*Significant at 95%

Table 23

Analysis of Variance Table for Grade Level by Social Adjustment Rating

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	4	1.09	0.27	0.28	0 .8878
Within groups	319	305.61	0.96		
Total	323	306.7			
Group:	Count	Mean	Standard Deviation	Standard Error	
Grade 1	65	2.2	0.85	0.11	
Grade 2	80	2.3	0.95	0.11	
Grade 3	73	2.25	0.89	0.1	
Grade 4	39	2.18	1.1	0.18	
Grade 5	67	2.34	1.14	0.14	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
1 vs 2	-0.1	0.32	0.09	0.61	
1 vs 3	-0.05	0.33	0.02	0.28	
1 vs 4	0.02	0.39	0.00268	0.1	
1 vs 5	-0.14	0.34	0.18	0.84	
2 vs 3	0.05	0.31	0.03	0.34	
2 vs 4	0.12	0.38	0.1	0.63	
2 vs 5	-0.04	0.32	0.02	0.27	
3 vs 4	0.07	0.38	0.03	0.35	
3 vs 5	-0.1	0.33	0.09	0.58	
4 vs 5	-0.16	0.39	0.17	0.83	

Table 24

Analysis of Variance Table for Grade Level by Portfolio Rating

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	4	152.34	38.08	32.42	0.0001
Within groups	256	300.68	1.17		
Total	260	453.02			
Group:	Count	Mean	Standard Deviation	Standard Error	
Grade 1	68	4.13	1.22	0.15	
Grade 2	75	2.55	0.84	0.1	
Grade 3	29	3.24	0.69	0.13	
Grade 4	28	2.21	0.99	0.19	
Grade 5	61	2.28	1.34	0.17	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
1 vs 2	1.59	0.36*	19.09*	8.74	
1 vs 3	0.89	0.47*	3.44*	3.71	
1 vs 4	1.92	0.48*	15.53*	7.88	
1 vs 5	1.85	0.38*	23.52*	9.7	
2 vs 3	-0.69	0.47*	2.15	2.93	
2 vs 4	0.33	0.47	0.48	1.38	
2 vs 5	0.27	0.37	0.51	1.43	
3 vs 4	1.03	0.57*	3.2*	3.58	
3 vs 5	0.96	0.48*	3.88*	3.94	
4 vs 5	-0.06	0.49	0.02	0.26	

*Significant at 95%

Table 25

Analysis of Variance Table for Grade Level by Reading Comprehension Achievement

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	4	9574.03	2393.51	8.33	0.0001
Within groups	315	90544.9	287.44		
Total	319	100118.92			
Group:	Count	Mean	Standard Deviation	Standard Error	
Grade 1	69	72.28	13.95	1.68	
Grade 2	69	56.33	18.78	2.26	
Grade 3	68	62.78	18.75	2.27	
Grade 4	60	59.93	16.86	2.18	
Grade 5	54	63.26	15.65	2.13	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
1 vs 2	15.94	5.68*	7.63*	5.52	
1 vs 3	9.5	5.7*	2.69*	3.28	
1 vs 4	12.34	5.89*	4.25*	4.12	
1 vs 5	9.02	6.06*	2.14	2.93	
2 vs 3	-6.45	5.7*	1.24	2.23	
2 vs 4	-3.6	5.89	0.36	1.2	
2 vs 5	-6.93	6.06*	1.26	2.25	
3 vs 4	2.85	5.91	0.22	0.95	
3 vs 5	-0.48	6.08	0.01	0.16	
4 vs 5	-3.33	6.26	0.27	1.05	

*Significant at 95%

Table 26

Analysis of Variance Table for Grade Level by Language Arts Achievement

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	4	4430.8	1107.7	3.16	0 .0144
Within groups	315	110393.09	350.45		
Total	319	114823.89			
Group:	Count	Mean	Standard Deviation	Standard Error	
Grade 1	69	64.81	25.33	3.05	
Grade 2	69	61.35	19.32	2.33	
Grade 3	68	72.13	15.14	1.84	
Grade 4	60	63.68	15.57	2.01	
Grade 5	54	64.87	14.86	2.02	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
1 vs 2	3.46	6.27	0.3	1.09	
1 vs 3	-7.32	6.29*	1.31	2.29	
1 vs 4	1.13	6.5	0.03	0.34	
1 vs 5	-0.06	6.69	0.0000747	0.02	
2 vs 3	-10.78	6.29*	2.84*	3.37	
2 vs 4	-2.34	6.5	0.12	0.71	
2 vs 5	-3.52	6.69	0.27	1.04	
3 vs 4	8.45	6.52*	1.62	2.55	
3 vs 5	7.26	6.71*	1.13	2.13	
4 vs 5	-1.19	6.91	0.03	0.34	

*Significant at 95%

Table 27

Analysis of Variance Table for Grade Level by Mathematics Achievement

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	4	7704.39	1926.1	5.59	0.0002
Within groups	319	109919.71	344.58		
Total	323	117624.1			
Group:	Count	Mean	Standard Deviation	Standard Error	
Grade 1	68	65.34	18.88	2.29	
Grade 2	70	59.79	18.87	2.26	
Grade 3	69	68.36	19.46	2.34	
Grade 4	61	54.43	16.58	2.12	
Grade 5	56	59.55	18.7	2.5	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
1 vs 2	5.55	6.22	0.77	1.76	
1 vs 3	-3.02	6.24	0.23	0.95	
1 vs 4	10.91	6.44*	2.78*	3.33	
1 vs 5	5.78	6.59	0.75	1.73	
2 vs 3	-8.58	6.2*	1.85	2.72	
2 vs 4	5.36	6.4	0.68	1.65	
2 vs 5	0.23	6.55	0.00122	0.07	
3 vs 4	13.94	6.42*	4.56*	4.27	
3 vs 5	8.81	6.57*	1.74	2.64	
4 vs 5	-5.13	6.76	0.56	1.49	

*Significant at 95%

Table 28

Analysis of Variance Table for Grade Level by Total Battery Achievement

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	4	4247.82	1061.96	4.03	0 .0033
Within groups	314	82760.27	263.57		
Total	318	87008.09			
Group:	Count	Mean	Standard Deviation	Standard Error	
Grade 1	69	64.22	16.12	1.94	
Grade 2	68	57.15	16.45	1.99	
Grade 3	68	65.88	18.14	2.2	
Grade 4	59	57.71	15.08	1.96	
Grade 5	55	63.95	14.72	1.99	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
1 vs 2	7.07	5.46*	1.62	2.55	
1 vs 3	-1.66	5.46	0.09	0.6	
1 vs 4	6.51	5.66*	1.28	2.26	
1 vs 5	0.27	5.77	0.00215	0.09	
2 vs 3	-8.74	5.48*	2.46*	3.14	
2 vs 4	-0.56	5.68	0.01	0.2	
2 vs 5	-6.8	5.79*	1.33	2.31	
3 vs 4	8.17	5.68*	2	2.83	
3 vs 5	1.94	5.79	0.11	0.66	
4 vs 5	-6.23	5.99*	1.05	2.05	

*Significant at 95%

Table 29

Analysis of Variance Table for Special Education Participation by Academic Rating

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	73.46	73.46	80.53	0 .0001
Within groups	329	300.14	0.91		
Total	330	373.6			
Group:	Count	Mean	Standard Deviation	Standard Error	
Participants	30	3.93	0.87	0.16	
NonParticipants	301	2.29	0.96	0.06	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Part vs. NonPar	1.64	0.36*	80.53*	8.97	

*Significant at 95%

Table 30

Analysis of Variance Table for Special Education Participation by Social Adjustment Rating

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	17.33	17.33	19.29	0 .0001
Within groups	322	289.37	0.9		
Total	323	306.7			
Group:	Count	Mean	Standard Deviation	Standard Error	
Participants	29	3	1.2	0.22	
NonParticipants	295	2.19	0.92	0.05	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Part vs. NonPar	0.81	0.36*	19.29*	4.39	

*Significant at 95%

Table 31

Analysis of Variance Table for Special Education Participation by Portfolio Rating

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	14.38	14.38	8.49	0 .0039
Within groups	259	438.64	1.69		
Total	260	453.02			
Group:	Count	Mean	Standard Deviation	Standard Error	
Participants	27	3.63	1.11	0.21	
NonParticipants	234	2.86	1.32	0.09	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Part vs. NonPar	0.77	0.52*	8.49*	2.91	

*Significant at 95%

Table 32

Analysis of Variance Table for Special Education Participation by Reading Comprehension Achievement

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	5103.52	5103.52	17.08	0 .0001
Within groups	318	95015.4	298.79		
Total	319	100118.92			
Group:	Count	Mean	Standard Deviation	Standard Error	
Participants	11	41.82	19.78	5.96	
NonParticipants	309	63.74	17.2	0.98	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Part vs. NonPar	-21.92	10.44*	17.08*	4.13	

*Significant at 95%

Table 33

Analysis of Variance Table for Special Education Participation by Language Arts Achievement

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	1110.47	1110.47	3.11	0 .079
Within groups	318	113713.41	357.59		
Total	319	114823.89			
Group:	Count	Mean	Standard Deviation	Standard Error	
Participants	11	55.55	17.59	5.3	
NonParticipants	309	65.77	18.95	1.08	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Part vs. NonPar	-10.22	11.42	3.11	1.76	

Table 34

Analysis of Variance Table for Special Education Participation by Mathematics Achievement

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	2873.71	2873.71	8.06	0 .0048
Within groups	322	114750.39	356.37		
Total	323	117624.1			
Group:	Count	Mean	Standard Deviation	Standard Error	
Participants	14	47.71	15.48	4.14	
NonParticipants	310	62.36	19.01	1.08	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Part vs. NonPar	-14.65	10.15*	8.06*	2.84	

*Significant at 95%

Table 35

Analysis of Variance Table for Special Education Participation by Total Battery Achievement

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	3702.46	3702.46	14.09	0 .0002
Within groups	317	83305.63	262.79		
Total	318	87008.09			
Group:	Count	Mean	Standard Deviation	Standard Error	
Participants	12	44.58	19.49	5.63	
NonParticipants	307	62.49	16.08	0.92	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Part vs. NonPar	-17.91	9.39*	14.09*	3.75	

*Significant at 95%

Table 36

Analysis of Variance Table for Chapter 1/L.A.P. Participation by Academic Rating

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	50.01	50.01	50.84	0 .0001
Within groups	329	323.6	0.98		
Total	330	373.6			
Group:	Count	Mean	Standard Deviation	Standard Error	
Participants	35	3.57	0.85	0.14	
NonParticipants	296	2.31	1.01	0.06	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Part. vs. NonPart.	1.26	0.35*	50.84*	7.13	

*Significant at 95%

Table 37

Analysis of Variance Table for Chapter 1/L.A.P. Participation by Social Adjustment Rating

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	14.72	14.72	16.24	0 .0001
Within groups	322	291.98	0.91		
Total	323	306.7			
Group:	Count	Mean	Standard Deviation	Standard Error	
Participants	32	2.91	1.06	0.19	
NonParticipants	292	2.19	0.94	0.06	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Part. vs. NonPart.	0.71	0.35*	16.24*	4.03	

*Significant at 95%

Table 38

Analysis of Variance Table for Chapter 1/L.A.P. Participation by Portfolio Rating

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	13.08	13.08	7.7	0 .0059
Within groups	259	439.94	1.7		
Total	260	453.02			
Group:	Count	Mean	Standard Deviation	Standard Error	
Participants	31	3.55	1.09	0.2	
NonParticipants	230	2.86	1.33	0.09	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Part. vs. NonPart.	0.69	0.49*	7.7*	2.77	

*Significant at 95%

Table 39

Analysis of Variance Table for Chapter 1/L.A.P. Participation by Reading Comprehension Achievement

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	7542.08	7542.08	25.91	0 .0001
Within groups	318	92576.84	291.12		
Total	319	100118.92			
Group:	Count	Mean	Standard Deviation	Standard Error	
Participants	31	48.16	21.04	3.78	
NonParticipants	289	64.57	16.59	0.98	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Part. vs. NonPart.	-16.41	6.34*	25.91*	5.09	

*Significant at 95%

Table 40

Analysis of Variance Table for Chapter 1/L.A.P. Participation by Language Arts Achievement

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	9326.11	9326.11	28.11	0 .0001
Within groups	318	105497.78	331.75		
Total	319	114823.89			
Group:	Count	Mean	Standard Deviation	Standard Error	
Participants	31	48.94	25.53	4.58	
NonParticipants	289	67.19	17.28	1.02	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Part. vs. NonPart.	-18.25	6.77*	28.11*	5.3	

*Significant at 95%

Table 41

Analysis of Variance Table for Chapter 1/L.A.P. Participation by Mathematics Achievement

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	4235.42	4235.42	12.03	0 .0006
Within groups	322	113388.68	352.14		
Total	323	117624.1			
Group:	Count	Mean	Standard Deviation	Standard Error	
Participants	31	50.61	23.56	4.23	
NonParticipants	293	62.9	18.2	1.06	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Part. vs. NonPart.	-12.29	6.97*	12.03*	3.47	

*Significant at 95%

Table 42

Analysis of Variance Table for Chapter 1/L.A.P. Participation by Total Battery Achievement

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	6738.28	6738.28	26.61	0 .0001
Within groups	317	80269.81	253.22		
Total	318	87008.09			
Group:	Count	Mean	Standard Deviation	Standard Error	
Participants	31	47.81	21.9	3.93	
NonParticipants	288	63.32	15.15	0.89	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Part. vs. NonPart.	-15.52	5.92*	26.61*	5.16	

*Significant at 95%

Table 43

Analysis of Variance Table for Gifted Program Participation by Academic Rating

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	27.1	27.1	25.73	0 .0001
Within groups	329	346.5	1.05		
Total	330	373.6			
Group:	Count	Mean	Standard Deviation	Standard Error	
Participants	25	1.44	0.87	0.17	
NonParticipants	306	2.52	1.04	0.06	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Part. vs. NonPart.	-1.08	0.42*	25.73*	5.07	

*Significant at 95%

Table 44

Analysis of Variance Table for Gifted Program Participation by Social Adjustment Rating

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	1.86	1.86	1.97	0.1615
Within groups	322	304.84	0.95		
Total	323	306.7			
Group:	Count	Mean	Standard Deviation	Standard Error	
Participants	25	2	1.08	0.22	
NonParticipants	299	2.28	0.96	0.06	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Part. vs. NonPart.	-0.28	0.4	1.97	1.4	

Table 45

Analysis of Variance Table for Gifted Program Participation by Portfolio Rating

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	15.24	15.24	9.01	0 .0029
Within groups	259	437.78	1.69		
Total	260	453.02			
Group:	**Count**	**Mean**	**Standard Deviation**	**Standard Error**	
Participants	20	2.1	0.91	0.2	
NonParticipants	241	3.01	1.33	0.09	
Comparison:	**Mean Difference**	**Fisher PLSD**	**Scheffe F-test**	**Dunnett t**	
Part. vs. NonPart.	-0.91	0.6*	9.01*	3	

*Significant at 95%

Table 46

Analysis of Variance Table for Gifted Program Participation by Reading Comprehension Achievement

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	1749.11	1749.11	5.65	0 .018
Within groups	318	98369.81	309.34		
Total	319	100118.92			
Group:	Count	Mean	Standard Deviation	Standard Error	
Participants	26	70.85	16.96	3.33	
NonParticipants	294	62.29	17.64	1.03	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Part. vs. NonPart.	8.56	7.08*	5.65*	2.38	

*Significant at 95%

Table 47

Analysis of Variance Table for Gifted Program Participation by Language Arts Achievement

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	2577.07	2577.07	7.3	0 .0073
Within groups	318	112246.82	352.98		
Total	319	114823.89			
Group:	Count	Mean	Standard Deviation	Standard Error	
Participants	26	74.96	19.4	3.8	
NonParticipants	294	64.57	18.73	1.09	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Part. vs. NonPart.	10.39	7.56*	7.3*	2.7	

*Significant at 95%

Table 48

Analysis of Variance Table for Gifted Program Participation by Mathematics Achievement

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	3279.92	3279.92	9.24	0 .0026
Within groups	322	114344.18	355.11		
Total	323	117624.1			
Group:	Count	Mean	Standard Deviation	Standard Error	
Participants	26	72.5	17.41	3.41	
NonParticipants	298	60.79	18.96	1.1	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Part. vs. NonPart.	11.71	7.58*	9.24*	3.04	

*Significant at 95%

Table 49

Analysis of Variance Table for Chapter 1/L.A.P. Participation by Total Battery Achievement

Source:	DF	Sum Squares	Mean Square	F-test	P value
Between groups	1	6738.28	6738.28	26.61	0 .0001
Within groups	317	80269.81	253.22		
Total	318	87008.09			
Group:	Count	Mean	Standard Deviation	Standard Error	
Participants	31	47.81	21.9	3.93	
NonParticipants	288	63.32	15.15	0.89	
Comparison:	Mean Difference	Fisher PLSD	Scheffe F-test	Dunnett t	
Part. vs. NonPart.	-15.52	5.92*	26.61*	5.16	

*Significant at 95%